

~~CONFIDENTIAL~~

CODIB-D-29  
26 March 1959

UNITED STATES INTELLIGENCE BOARD  
COMMITTEE ON DOCUMENTATION

MEMORANDUM FOR: Assistant Director, CR

SUBJECT : Trip Report

I. Introduction

- a. The observations supplied herewith are based on one-day visits by myself and Theodore Wagman to each of the following organizations, 18 - 20 March 1959:

1. Documentation Center, ATIC, Wright-Patterson Air Force Base, Dayton, Ohio.
2. Project White Stork, Columbus, Ohio.
3. Air Gas Turbine Library System, Air Gas Turbine Division, General Electric, Cincinnati, Ohio.

- b. The visit to ATIC was concerned primarily with their planning for SOV-STEP (Soviet Scientific and Technical Exploitation Program) which was of interest from the standpoint of design and management of an information system. Its substantive aspects are, of course, within the province of the USIB Committee on Exploitation.

c.

In addition, two of its personnel, Messrs G. S. Simpson Jr., and J. W. Murdock discussed with us a paper of CODIB interest entitled "A Concept of a Standard Intelligence Information Reporting System."

- d. The General Electric visit concerned their recent application of an IBM 704 computer to the operation of a coordinate indexing system for storage and retrieval of scientific abstracts. The discussions did not involve classified subjects.

~~CONFIDENTIAL~~

25X1

**Page Denied**

Next 1 Page(s) In Document Denied

C-O-N-F-I-D-E-N-T-I-A-L

- 4 -

CODIB-D-29  
26 March 1959

The twenty professional information staff are the key to the system. Each is assigned responsibility for one or several of the subject fields in which FWS is active. Each is picked for his academic and experience qualifications and his salary is set between \$5,000 and \$10,000 annually depending on these factors. He is regularly assigned to work with the engineers on projects falling within his subject field, travels with them and serves in general as an alter ego in matters concerning documentation. He indexes all information received at FWS which is pertinent to his subject assignment and is personally responsible for the arrangement and use of the indexing product. The information intake which is routed successively to all information specialists now averages:

<u>Category</u>	<u>Per Month</u>
USIB information reporting	2,000 documents
AF & SOV-STEP abstract cards	4,500 abstracts
AID Infa Div (LC) book abatracts	1,000 abstracts
Open literature extracts (e.g. N.Y. Times)	1,000 extracts

Indexing consists of the underlining of words in the text of the given document or the extracts therefrom which are selected for incorporation in FWS files. Five major files result:

Personality  
Facility  
Location  
Source  
Clue Word

The underlining of a word, personality, place name, etc. will cause a copy of the basic 5 x 8 entry card to be filed under that indicator. The underlining of all index terms on each card entry permits the user to pursue his subject on what may be termed a free association basis wherever his own thinking may lead him.

Because of security considerations the consumer comes to the file. There is no document storage or reproduction service required with the exception of a FWS policy of microcarding extracted documents so that the original may be consulted in full whenever the user finds this desirable.

C-O-N-F-I-D-E-N-T-I-A-L

C-O-N-F-I-D-E-N-T-I-A-L

- 5 -

CODIB-D-29  
26 March 1959

The personality and facility files parallel in many respects the [redacted] systems in OCR. The clue word file, however, is the critical file from the point of view of a subject retrieval system. Messrs Simpson and Murdock who are senior officials at the Project emphasized from the outset that their system rejects subject indexing as a philosophy on the theory that classification freezes information into rigid categories thereby tending to conceal it from future searchers employing new language for new interests. Each document therefore provides its own index language. No classification language may be imposed. We did find, however, that many elements of classification enter into the system in the course of arrangement of the clue word files by the information specialist. Guide cards in the files proceeded from the general to the specific, showed chronological order (based on the date the information became known to FWS) and grouping of subjects by area.

25X1

The subject interests of FWS were described in terms of a pyramid with a broad base of basic sciences, moving through applied sciences to AF projects and culminating in weapons data at the apex. The availability of intelligence parallels this sequence with the most plentiful being that concerning the basic sciences, e.g. Soviet scientific literature.

25X1

Since categorization, if not classification, is a useful tool for the human mind, the clue word system raises many questions relating to focus in literature searching. Obviously the information specialist is essential as the catalyst who must solve problems of cross reference, synonyms, redundancy, ambiguity and conflicting data. It was indicated incidentally that a pruning of some 25% of the 8-year old files is now believed appropriate. This is being established by segregating segments of the file and monitoring their use.

In the course of the discussions, we also dealt with the problem of standardizing USIB information reports in terms of their format, editorial features, dissemination and security control. Messrs Simpson and Murdock had written in April 1958, a paper entitled "A Concept of a Standard Intelligence Information Reporting System" on which we commented in terms of CODIB programs and of current proposals for punched tape transmission of reports from field to headquarters, through automatic dissemination systems and possibly into consumer offices.

C-O-N-F-I-D-E-N-T-I-A-L

~~O-N-F-I-D-E-N-T-I-A-L~~

- 6 -

CODIB-D-29  
26 March 1959

#### CONCLUSIONS:

PWS personnel should be invited to brief CODIB on their system and it might be profitable to schedule a CODIB visit to the installation.

Following the establishment of closer liaison between PWS and the CODIB community, it would be desirable to obtain their operating manuals and to promote widespread understanding of the advantages and limitations of the clue word approach.

The paper by Messrs Simpson and Mardock should be requested of AFCEC-4 for distribution as a CODIB document.

- c. Air Gas Turbine Library System, Air Gas Turbine Division, General Electric:

Early in 1958, the GE librarians secured the assistance of their computer personnel for the project of transferring their manual coordinate indexing system to an IBM 704 computer. Three programs were written for the computer of approximately 2,000 instructions each, as follows:

1. The index posted to magnetic tape and consisting of keywords with document numbers arranged in sequence under each keyword. There are 9,000 terms in the keyword dictionary.
2. The store of scientific abstracts (covered by the keyword index) posted in document number order to magnetic tapes at 10,000 abstracts per tape. Approximately 38,000 abstracts are now posted and the file is growing at the rate of 200 items per week.
3. The routines for updating the index and abstract files. Updating is performed simultaneously with literature searching. Up to 99 searches may be handled at one time.

The GE personnel indicated that the project is in a very early stage of development. The programming only provides for matching a logical sum - at present. The ability to organize searches and search products by date, source, and author remain to be established. It was indicated that the National Science Foundation is being solicited for support of the further development of the System.

As at Project White Stork, GE employs information specialists with engineering backgrounds to work with the company's engineers in conducting literature searches.

25X1

- 7 -

CODIB-D-29  
26 March 1959

The Technical Library Manager, Mr. Bernard Dennis, demonstrated to us a Bell Telephone conversation recorder (rental \$30 monthly). AGT engineers are being supplied with the appropriate telephone number and may henceforth call in at night from their homes and read in extracts from their reading together with suggestions for key-words under which the information should be indexed. The information will be processed into the computer thus enabling an engineer to recover his own input at a later date and to conduct searches according to terms which he has personally recommended.

**CONCLUSIONS:**

At its present stage, this application adds nothing to the theory of coordinate indexing. It constitutes a very minor activity of the GE computer center and could in no sense justify itself at present as the principal application of such equipment. Attempts to develop the programming will be followed with interest.

STAT

Special Assistant to the AD  
Central Reference

C-O-N-F-I-D-E-N-T-I-A-L